

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204 Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, D.C. 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 11 May 1998	3. REPORT TYPE AND DATES COVERED Final technical report (9/30/92 - 9/30/96)	
4. TITLE AND SUBTITLE Establishment of Acoustic Reverberation Special Research Program (ARSRP) Data Archive			5. FUNDING NUMBERS N00014-92-J-4079	
6. AUTHOR(S) John A. Orcutt				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of California, San Diego Scripps Institution of Oceanography IGPP (0225) La Jolla, CA 92093			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research Ocean Acoustics 800 North Quincy Street Arlington, VA 22217-5000			10. SPONSORING /MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES			19980514 154	
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				
13. ABSTRACT (Maximum 200 words) We have developed a database containing digital data collected during the Office of Naval Research Bottom/Subbottom Acoustic Reverberation Special Research Program (ONR ARSRP). The data entered into the database include the acoustic data from the 1991 Acoustic Reverberation Special Research Program, the bottom bathymetry, side-scan data, and seismic reflection data from the 1992 Large-Scale Geophysics Cruise, the digital data collected during the 1993 Small-Scale Geophysics Cruise, and the acoustic data collected on the R/V Cory Chouest, the R/V Knorr, and the R/V Alliance during the 1993 Acoustics Experiment. The data is accessible by scientific investigators through the Internet or through copies to Exabyte and/or DAT tapes or magneto-optical disks at IGPP. The system provides an essentially permanent archive of the ARSRP data which as been collected at great expense - and encourages the thorough analysis of this acoustic data.				
14. SUBJECT TERMS Data archive, acoustic data			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT unclassified	20. LIMITATION OF ABSTRACT SAR	

Final Report:
**Establishment of Acoustic Reverberation Special Research
Program (ARSRP) Data Archive**

UCSD 93-1046 (ONR N00014-92-J-4079)

Professor John A. Orcutt
521-60-8566
(619)534-2887
jorcutt@igpp.ucsd.edu

Original Abstract of Proposed Research:

We propose to develop a database containing digital data collected during the Office of Naval Research Bottom/Subbottom Acoustic Reverberation Special Research Program (ONR ARSRP). The data to be entered into the database include the acoustic data from the 1991 Acoustic Reverberation Special Research Program, the bottom bathymetry, side-scan data, and seismic reflection data from the 1992 Large-Scale Geophysics Cruise, the digital data to be collected during the 1993 Small-Scale Geophysics Cruise, and the acoustic data to be collected on the R/V Cory Chouest, the R/V Knorr, and the R/V Alliance during the 1993 Acoustics Experiment. The archive medium will be 12" optical Write-Once, Read-Many (WORM) disks mounted in an Epoch optical jukebox. The data will be accessible by investigators through the Internet or through copies to Exabyte or DAT (Digital Audio Tape) tapes or magneto-optical disks at IGPP. On-line access will be available through the period of the ARSRP and until such time as funding requires retiring the disks to an archive for long-term data storage. Very importantly, the system will provide an essentially permanent archive of the ARSRP data which have been and will be collected at great expense.

Long Range Objectives:

Long-term access to acoustic data from the ARSRP field programs will be archived at the Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics for use by investigators at IGPP and throughout the scientific community.

1. Introduction

The ONR ARSRP program began to accumulate large quantities of data in 1991 with the completion of the Reconnaissance Experiment in the ONR Atlantic Natural Laboratory. This experiment produced 384 GB of acoustic data and the task of making the data readily available for analysis by the community became a major problem. The Office of Naval Research provided an interesting solution to this problem through the funding of this grant.

We originally purchased an RSS-48 mass store from Metrum Technologies. Originally, the system was able to store about 660 GB of data and this provided the storage space needed to make the acoustic data available to ARSRP investigators. This device uses a VHS-format tape system, a

successor to the VLDS drives in wide use in the Navy, to store digital data on 48 cassettes. The cassettes are fed to two separate drives by a robot mechanism which is driven by AMASS software which emulates a commercial hierarchical file system modeled on that used by the SUN operating system. The system was originally front-ended by a SUN clone workstation and included 1 GB of disk storage which served as cache and contained the database needed for accessing the files and 2 GB which provide space for user-controlled staging. The entire mass store and database are continuously backed up while operating in the background during low usage periods. After several months of operation, the capacity of the media was increased so the mass store could provide on-line access to 858 GB of data. The capacity of the system was once-again increased to a full TByte. When the physical capacity of the mass store is exceeded, the data sets can be taken off line and stored on the shelf. The access information for the data remains in the on-line database so users can still obtain access. The operator is notified to reinstall an off-line tape in this instance.

2. Evolution of the Current Mass Store

Early in 1995 IGPP was funded by the National Science Foundation, with a contribution of matching

funds from IGPP and Scripps, to add a second Metrum to the mass store and to upgrade the SUN clone to a much higher-capacity Silicon Graphics system. At that time, the configuration of the mass store is shown in Figure 1. The primary function of the mass store remained that of providing data access to ARSRP investigators. However, the utility of the system had grown to the point that other investigators were beginning to store digital data. Users from outside Scripps were growing in number and interests.

In response to this interest, we have again obtained funding from the National Science Foundation to upgrade the mass store. The current mass store at IGPP is shown in Figure 2. While both Metrums continue to be used, the heart of the system has become a StorageTek 9710 Library Storage Module - a robot built around six Digital Linear Tape (DLT) tape drives. The capacity of the new mass store is now 10 TB and is easily expandable to twice that

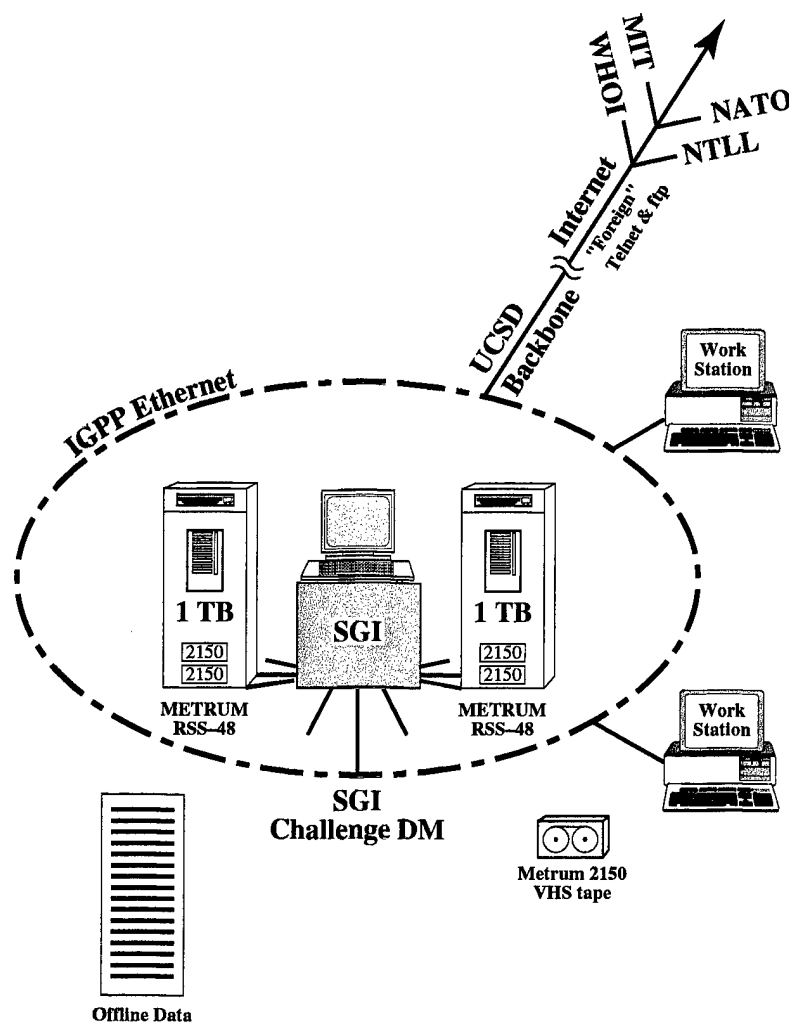


Figure 1: IGPP 2 TB Mass Store in the period 1995-97.

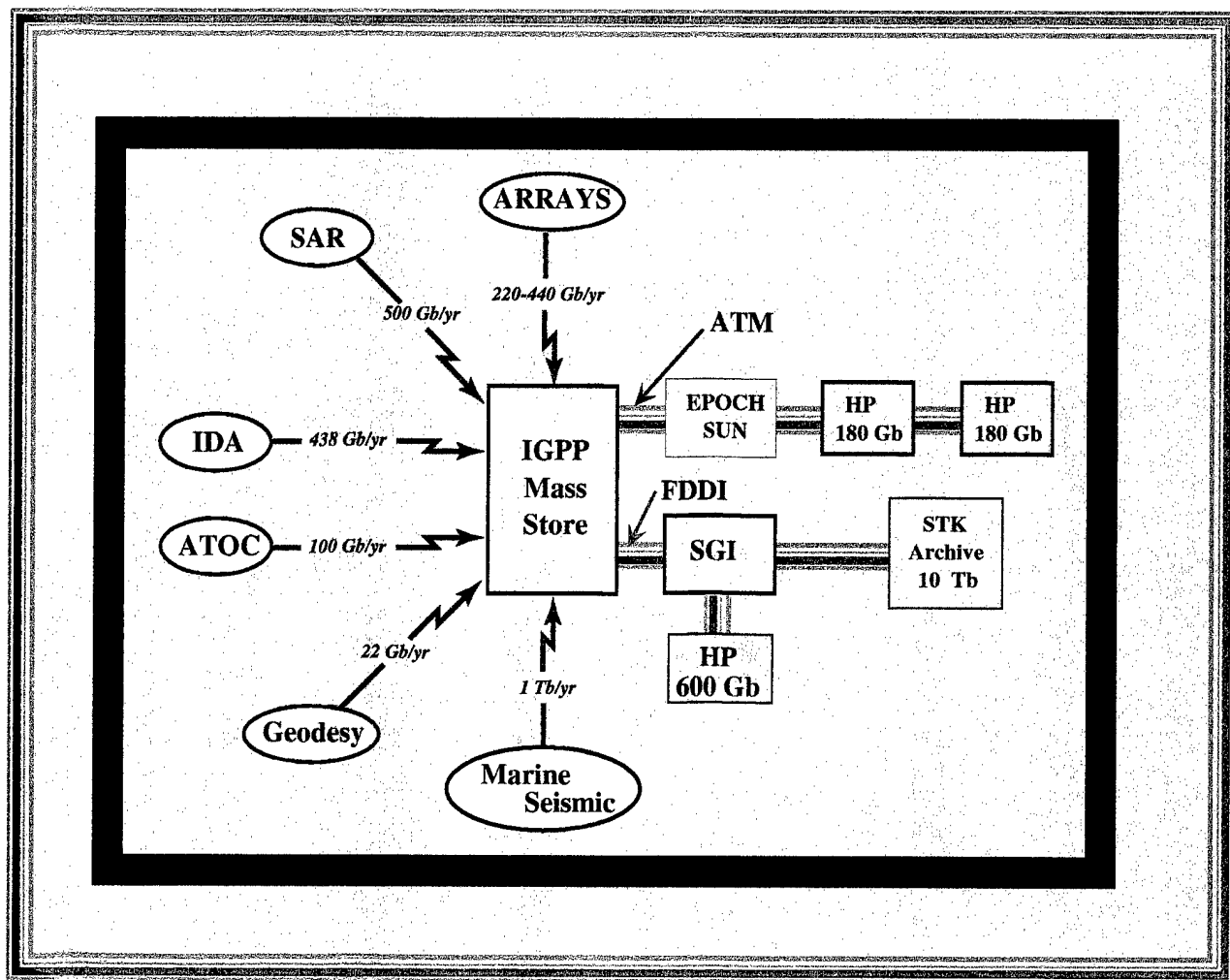


Figure 2: The current IGPP mass store. The mass store is connected to IGPP users and the outside world through an ATM (OC-3) network featuring data transfer at 155 mbps.

capacity. In addition, we've added a HP 600 GB magneto-optical disk robot to the system to provide random-access to large quantities of data.

During this month, one of the two Metrums will be removed from the mass store and surplused and the second will follow in June. All of the original ONR equipment will have been retired by mid-year although the data originally stored will continue to be available for remote access. Presently, 687,140 MB of ARSRP data are available and the total usage is 2.27 TB. The original ONR investment in the mass store has allowed us to leverage a very sophisticated mass store system which has become useful to a broad scientific community. In Figure 2, *marine seismic* refers to seafloor and multichannel seismic data (a recent 3-D experiment on the East Pacific Rise produced 0.5 TB), *Geodesy* is a database consisting of global GPS data which are used daily at IGPP by the Scripps Orbit and Permanent Array Center in computing monument positions and GPS orbits, *ATOC* is the Acoustic Thermometry of Ocean Climate project, *IDA* refers to data telemetered to IGPP daily from its global network of seismographs, *SAR* is Synthetic Aperture Radar data from the Scripps/IGPP x-band satellite antenna used for receiving data from ERS-1/2, and *Arrays* are IGPP seismic arrays in California, Colorado, and Kirghizia. IGPP and MPL have established a smaller mass store, but one

which uses identical software, at MPL Point Loma for archiving classified data, particularly data from SOSUS arrays. While the hardware has evolved enormously and will continue to do so, the choice of a robust software/hardware combination has allowed simple transitions from system to system. Transporting data from the older Metrums to the new Storage Tek device required less than a month.

The cost of operating the system is quite reasonable. The maintenance costs are shown below:

\$12,840	Hardware
8,928	Software
8,193	SGI dual-processor
\$29,961	Total hardware maintenance
\$26,428	3 months of Paul Henkart time
\$56,389	Total Direct Costs
\$28,758	Indirect Costs
\$85,147	Total

The cost of maintaining a GB of data on-line for a year is \$8.51. The number will continue to decline as the mass store grows in size in coming years. I feel that it would be very important in the future to archive ONR ocean acoustic experiment data to ensure long-term access and to encourage the thorough analysis of the expensive data sets collected in the field. I would be very interested in discussing this with ONR.